Detecting House Numbers in Street View Imagery using Convolutional Neural Networks

Matthew Peyrard

November 14, 2016

1 Definition

1.1 Project Overview

This project is my submission for the Capstone Project section for Udacity's Machine Learning Nanodegree program. The goal of the project is to automate the task of identifying house numbers from imagery taken from Google's street view cameras.

1.2 Problem Statement

The process of *manually* identifying and cataloging such imagery is very expensive due to the scale at which such a process must be applied. However, if computational resources could be applied to the task with reasonably high accuracy, then the process could be sped up enormously, in addition to the costs saved from not having to hire thousands of people to perform the task manually. The solution was based on the original paper published by Goodfellow et al. [2], utilizing deep convolutional neural networks. The standard street view house number dataset [1] was used as a training source for this project.

1.3 Metrics

Performance for this task is done using an accuracy calculation based on a train/test split. Every 100 iterations of the training algorithm, the test set

is fed into the algorithm and we measure the accuracy. No partial credit is provided for the algorithm. A prediction is correct if and only if the entire digit is predicted correctly. Furthermore, since we are using the method described in the original paper [2], the predicted length of the digit must also be correct.

2 Analysis

2.1 Data Exploration

2.2 Exploratory Visualization

2.3 Algorithms and Techniques

2.4 Benchmark

3 Methodology

3.1 Data Preprocessing

3.2 Implementation

3.3 Refinement

4 Results

4.1 Model Evaluation and Validation

4.2 Justification

5 Conclusion

5.1 Free-Form Visualization

5.2 Reflection

5.3 Improvement

References

[1] Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, Andrew Y. Ng *Reading Digits in Natural Images with Unsupervised Feature Learning NIPS Workshop on Deep Learning and Unsupervised Feature Learning 2011.* http://ufldl.stanford.edu/housenumbers/

[2] Goodfellow, Ian J.; Bulatov, Yaroslav; Ibarz, Julian; Arnoud, Sacha; Shet, Vinay *Multi-digit Number Recognition from Street View Imagery using Deep Convolutional Neural Networks*